IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 8, May 2025
Impact Factor: 7.67

AI/ML-Assisted Cryptanalysis: A Deep Learning Approach to Cipher Pattern Detection

Prof. Pooja Kadam, Karamjeet Singh, Devashree Bendwar, Siddhant Sanjeev Badardini, Harsh Gupta
MIT ADT University, Loni-Kalbhor, Pune, Maharashtra, India

Abstract: *This study investigates the use of deep learning models Fully Connected Neural Networks (FCNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN)to detect patterns in encrypted data. Through 11 trials, we demonstrate that ML can identify plaintext types and weak ciphers (e.g., SHIFT, XOR) but fails against robust ciphers like AES, DES, and One-Time Pad (OTP) when properly keyed. The findings underscore MLs potential to augment cryptanalysis and the need for AI-aware encryption designs to counter evolving threats.*

Keywords: Cryptanalysis, Machine Learning, Neural Networks, AES, DES, SHIFT, XOR, OTP

I. INTRODUCTION

Cryptography secures digital communication through mathematical rigor, but the advent of artificial intelli- gence (AI) and machine learning (ML) has opened new avenues for automating cryptanalysis. This paper explores whether ML models can extract information from ciphertexts without key knowledge, specifically aiming to:

- Identify the encryption algorithm used (e.g., SHIFT, AES).
- Determine the plaintext type (e.g., English text, binary data).

Using the ML Crypto Analyzer framework, we conducted experiments with FCNN, CNN, and RNN models on ciphers including SHIFT, XOR, OTP, AES, and DES, evaluating their effectiveness in detecting cryptographic patterns.

II. BACKGROUND

Classical Cryptography

Cryptography transforms plaintext into ciphertext using algorithms and keys to ensure confidentiality and in- tegrity. Classical ciphers, such as the Caesar cipher (shifting letters, e.g., 'A' to 'D' with a shift of 3) and XOR cipher (byte-wise XOR with a key), are simple but vulnerable to frequency analysis and brute-force attacks due to small key spaces. Modern block ciphers, like AES (128/192/256-bit keys, 1014 rounds) and DES (56-bit keys, now outdated), employ complex operations for enhanced security, relying on computational hardness assumptions (e.g., factoring large integers).

Benefits of Traditional Cryptography:

- One-Time Pad (OTP) is unbreakable with truly random, one-time keys.
- Simple ciphers are easily implemented without computational resources.
- Effective against casual eavesdropping.

Drawbacks:

- Susceptible to frequency analysis and brute-force attacks.
- Key management in symmetric systems is challenging, risking interception.
- Limited scalability for large-scale or modern data formats.
- Lacks mechanisms for data integrity or authentication.

ML in Cryptanalysis

ML excels at detecting statistical anomalies, identifying:

- Biases in pseudorandom number generators.
- Correlations between plaintext and ciphertext features.

DOI: 10.48175/IJARSCT-26995

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 8, May 2025
Impact Factor: 7.67

- Anomalies in frequency distributions or temporal dependencies.

These capabilities challenge the cryptographic assumption that ciphertexts appear random. Prior works by LeCun et al. (2015) and Yang et al. (2021) demonstrate MLs ability to uncover subtle patterns, making it a powerful tool for cryptanalysis.

III. METHODOLOGY

System Architecture

The ML Crypto Analyzer framework consists of:

- **Data Generator:** Produces encrypted data using Python implementations of SHIFT, XOR, OTP, AES, and DES.
- **Feature Processor:** Converts plaintext/ciphertext to machine-readable formats (byte arrays, one-hot vectors).
- **ML Models:** FCNN (multi-layer perceptron), CNN (Conv1D for local patterns), RNN (LSTM for se- quential dependencies) in TensorFlow/Keras.
- **Evaluator:** Computes accuracy, loss, and confusion matrices.
- **Visualization Module:** Uses matplotlib for training metrics and visualizations

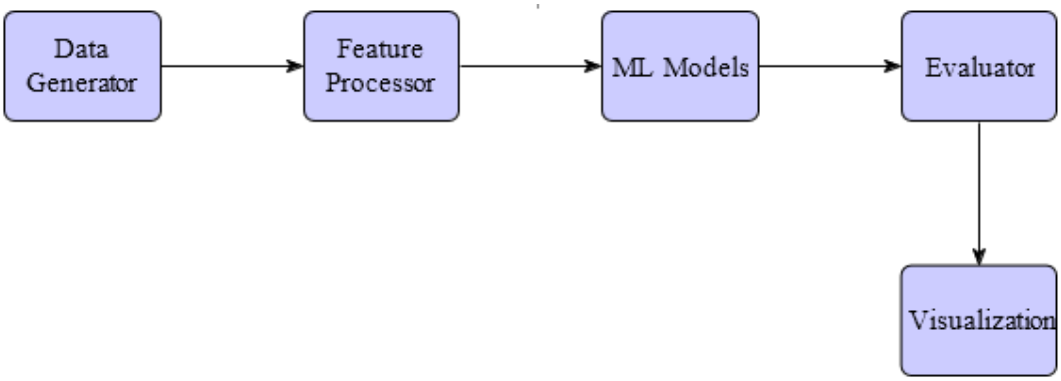

Figure 1: System Architecture

Process Flow

The process involves generating encrypted data, processing features, training models, evaluating performance, and visualizing results, with accuracy checks at each stage to ensure reliability

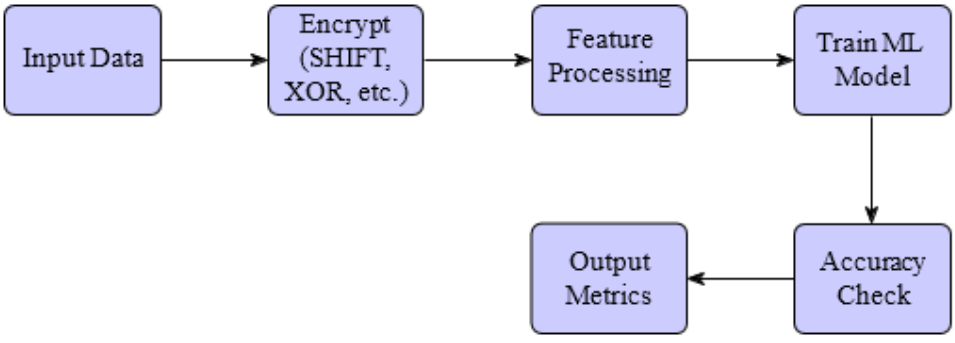

Figure 2: Process Flow

Experimental Design

We generated datasets using:

- **SHIFT Cipher:** Caesar-like character shifts.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 8, May 2025
Impact Factor: 7.67

- **XOR Cipher:** Byte-wise XOR with fixed or random keys.
- **OTP:** XOR with a truly random, one-time key
- **AES/DES:** Block ciphers with standard implementations.

Models were trained on synthetic datasets (batch sizes of 100 or 1000, 501000 epochs) to prevent overfitting, with performance evaluated using accuracy, loss, and confusion matrices.

Trials

Eleven trials tested:

- **Trial 1 (Sanity Test):** FCNN distinguished English text vs. binary data (high accuracy).
- **Trial 2 (SHIFT):** FCNN detected plaintext type, as SHIFT preserves frequency patterns.
- **Trial 3 (XOR):** FCNN identified plaintext type with fixed keys (success).
- **Trial 4 (OTP):** FCNN failed, validating OTPs perfect secrecy.
- **Trial 5 (XOR vs. SHIFT):** FCNN distinguished cipher types (high accuracy).
- **Trial 67 (AES/DES):** FCNN/CNN/RNN failed to detect plaintext type.
- **Trial 8 (AES vs. DES):** Partial success (80% accuracy) in distinguishing ciphers.
- **Trial 9 (AES vs. DES, Random Keys):** Failure, confirming random keys enhance security.
- **Trial 10 (SHIFT, Short Keys):** Success, showing weak obfuscation.
- **Trial 11 (SHIFT, Long Sequences):** RNN succeeded on 10,000-byte sequences.

IV. RESULTS

The trials revealed:

- **Weak Ciphers (SHIFT, XOR):** ML models detected plaintext types and cipher characteristics due to structural leakage, especially with fixed or short keys.
- **OTP:** Models failed, confirming its theoretical security.
- **AES/DES:** Resisted ML inference, particularly with random keys, though AES vs. DES classification achieved partial success (80% accuracy).

Figure 3: Results Summary

Trial	Cipher	Task	Result
1	None	Plaintext Type	Success
2	SHIFT	Plaintext Type	Success
3	XOR	Plaintext Type	Success
4	OTP	Plaintext Type	Failure
5	XOR/SHIFT	Cipher Type	Success
6	AES	Plaintext Type	Failure
7	DES	Plaintext Type	Failure
8	AES/DES	Cipher Type	Partial (80%)
9	AES/DES (Random Keys)	Cipher Type	Failure
10	SHIFT (Short Keys)	Plaintext Type	Success
11	SHIFT (Long)	Plaintext Type	Success

V. LIMITATIONS

- **Computational Constraints:** Limited hardware restricted model complexity, input sizes (up to 2000 bytes), and batch sizes.
- **Narrow Cipher Scope:** Only symmetric ciphers tested; asymmetric (RSA, ECC) and stream ciphers (ChaCha20) excluded.
- **Simplified Contexts:** Ideal conditions ignored real-world factors like padding or chaining modes.

DOI: 10.48175/IJARSCT-26995

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 8, May 2025
Impact Factor: 7.67

- **Weak Keys:** Some trials used fixed/weak keys, not reflecting best practices.
- **Model Diversity:** Advanced architectures (e.g., transformers) not explored.
- **Metrics:** Focused on accuracy, omitting precision, recall, or robustness analysis

VI. FUTURE WORK

Future research should:

- Test asymmetric ciphers (RSA, ECC) and stream ciphers (ChaCha20).
- Adopt advanced ML models (transformers, attention mechanisms).
- Simulate real-world protocols (TLS, SSH) with noise and multilayer encryption.
- Apply adversarial ML techniques (GANs, reinforcement learning).
- Expand datasets with multilingual text and multimedia.
- Enhance robustness and explainability with SHAP or LIME.
- Foster interdisciplinary collaboration for secure cipher design.

VII. CONCLUSION

This study demonstrates that ML models, particularly FCNNs and RNNs, can exploit weaknesses in simple ciphers (SHIFT, XOR) but fail against OTP and properly keyed AES/DES. The results highlight MLs potential to augment cryptanalysis while underscoring the robustness of modern ciphers. As ML evolves, encryption must adapt to maintain security, paving the way for AI-integrated cryptographic design.

REFERENCES

[1]. LeCun et al. (2015). Deep learning foundations for CNNs.
[2]. Krizhevsky et al. (2012). Advances in CNN architectures.
[3]. Yang et al. (2021). Unsupervised learning for nonlinear ciphers.
[4]. Roy Dar. ML Crypto Analyzer framework